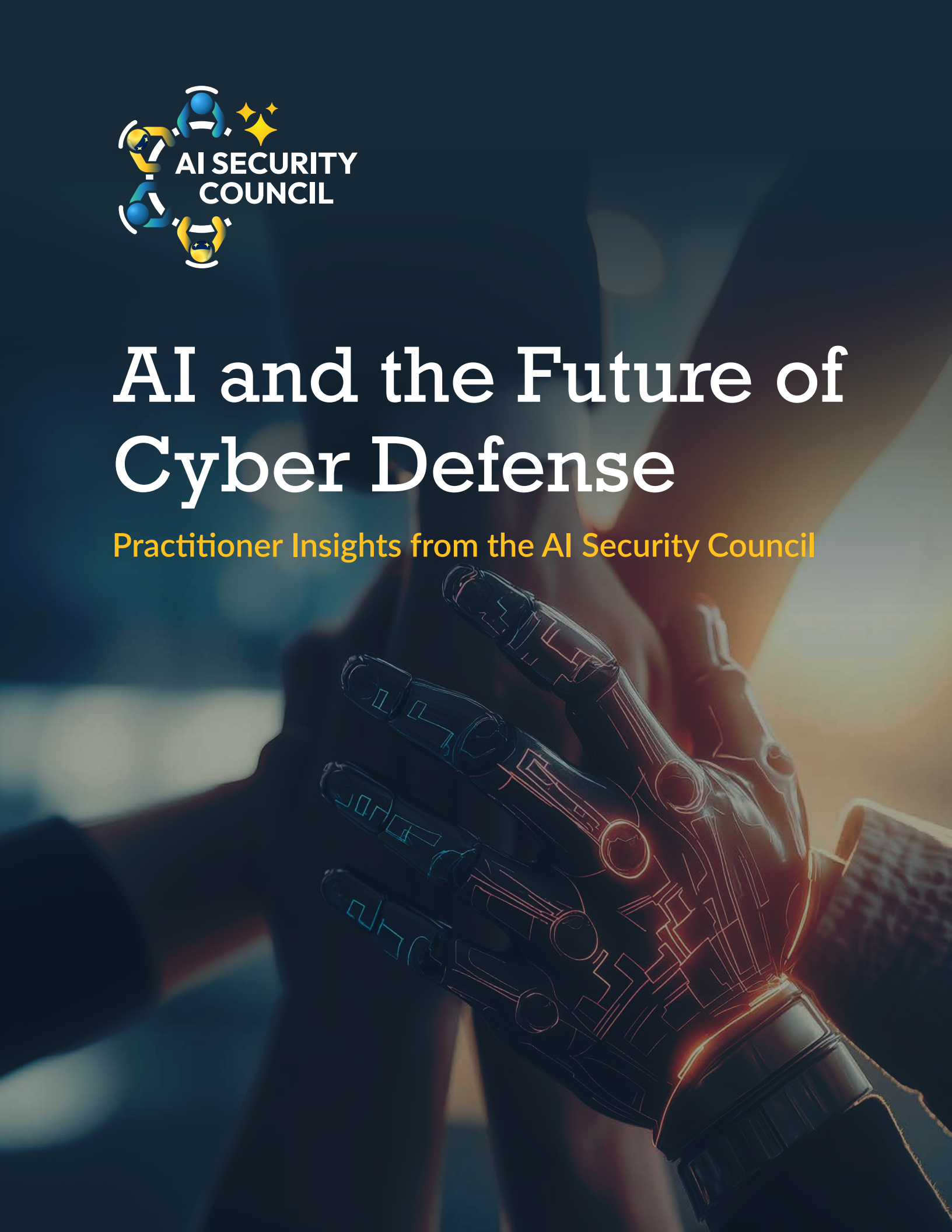
AI SECURITY COUNCIL

AI and the Future of Cyber Defense

Practitioner Insights from the AI Security Council



Artificial intelligence is collapsing the time and cost of cyber operations.

On the attacker's side, what used to take money, manpower, and specialized tools can now be done in seconds for next to nothing. Phishing and social engineering campaigns can now be generated in dozens of languages with a single prompt. Reconnaissance work, such as sifting through LinkedIn, GitHub, and exposed assets, uses AI crawlers that can stitch together a target's footprint in minutes. Malware authors are now turning to polymorphic code generation, constantly reshaping payloads to evade detection and defense mechanisms. Even people with little to no skill can act like seasoned operators, leveraging AI to write scripts or assemble attack playbooks that were previously out of reach. And when it comes to overwhelming defenders, the economics are even more lopsided, considering the bad guys can generate thousands of phishing variants or nearly free semi-legitimate probes, but every single one creates real costs on the defensive side.

Defenders, meanwhile, are trying to harness that same acceleration for themselves. AI is being folded into the SOC to compress triage that once took hours into minutes, helping analysts cut through mountains of low-value alerts. It enriches signals automatically with context from threat intelligence, asset data, and identity systems, providing analysts with a clearer picture before they even open the case. Teams are beginning to simulate attack paths with digital twins of their environment, validating whether a vulnerability actually leads to an exploitable breach rather than chasing every CVSS score. For lean security teams, AI is serving as a force multiplier by drafting incident reports, correlating logs, and even suggesting mitigation steps, allowing scarce human expertise to be applied where it matters most.

To separate hype from practice, the newly formed **AI Security Council** convened two private workshops in September 2025, bringing together 18 security leaders, including CISOs, engineers, and governance experts from enterprises, startups, and financial institutions.

Participating AI Security Council Members

AJ Debole
Field CISO
Oracle

Amy Lemberger
Senior Cybersecurity Consultant
Lemberger

Angelique Grado
Principal and CISO
Vervation LLC

Antoinette Stevens
Principal Security Engineer
Ramp

Aunudrei Oliver
Sr. Director IS
Allianz Life

Brandon Lindsay
Director of Information Security & Data Protection
HIAS

Cassandra Mack
CISO
Tensorwave

Chris DeNoia
CISO and Author

Korey Barrette
CISO
RSI Security

Jose Veitia
Director of Information Security

Sean Todd
CISO & AI Architect
Auditive

Susan Lloyd
Director of Information Governance & Security
iVisa

Ron Dilley
Principal Architect R&D
IS²

Ryan Rosado
Adjunct Faculty
Harvard

Adnan Dakhwe
CISO
DelphinusCyber

Sandip Wadje
Global Head of Emerging Tech Risks & Intelligence
BNP Paribas

Peter Holcomb
CEO
OptimoIT

Tomas Persson
CISO
Omegapoint

The discussions were structured around five guiding questions:

1. How is AI shifting the balance of power between attackers and defenders?
2. What does “effective defense” look like in the AI era?
3. How will the role of humans and AI in the SOC evolve?
4. What new risks, ethical, regulatory, and operational, emerge with autonomous defense, and how should CISOs govern them?
5. If advising a board today, where should enterprises prioritize investment to prepare for AI-driven threats and defenses?

Each participant shared written perspectives and then debated them live with their peers. What follows is a practitioner’s field guide, including lessons, patterns, and risks voiced directly by the people running security programs in this new world of AI in which we find ourselves.

How AI is Shifting the Balance of Power

Artificial intelligence is rewriting the economics of cyber conflict. On the offensive side, what once demanded money, manpower, and specialized tooling can now be achieved in seconds at almost no cost. Low-skill actors can generate multilingual phishing campaigns with a single prompt, utilize AI crawlers to stitch together target profiles in minutes, or leverage polymorphic code generation to produce malware that constantly shifts. The net result is that a lone attacker can suddenly operate with the reach of a seasoned team.

Council members described this shift as a collapse of barriers. **Chris DeNoia**, CISO and author, noted that: ***"AI is continuously lowering the bar… it allows a single actor to become ten times more effective than they were before."*** From a European vantage, **Tomas Persson**, CISO at Omegapoint, called it ***"democratization,"*** giving low-skilled hackers disproportionate leverage. **Korey Barrette**, CISO at RSI Security, pointed out that the barrier isn't only technical but cultural: ***"Many end users see AI as a black box. They'll push anything they can into it without understanding what's happening under the hood. Education is critical, or else the same lack of awareness that fuels phishing will fuel misuse of AI."***

For many, the imbalance is already visible. **Angelique Grado**, Principal and CISO, Vervation LLC, was direct: ***"Currently, the balance of power is with the attackers as AI has given them the advantage."*** Later in the workshop, she expanded this point through her '6A' framework for structuring AI defense. **Adnan Dakhwe**, CISO at DelphinusCyber, agreed: ***"attackers are moving faster today, even if defenders may catch up in time."***

Others, like **Ron Dilley**, Principal Architect R&D, IS² cautioned that the asymmetry isn't yet decisive: ***"I have not seen a material asymmetry now, but attackers benefit more from scale than we do. They only need to be right once."*** The asymmetry of risk, where attackers can fail cheaply but defenders pay dearly for every miss, remains a structural challenge.

The Council also flagged new dynamics that weren't just better-targeted attacks, but an overwhelming surge of noise. **Susan Lloyd**, Director of Information Governance at iVisa, described attackers ***"flooding environments with probes,"*** warning that defenders could drown in the fire hose. **Brandon Lindsay**, Director of Information Security at HIAS, added that AI-fueled phishing and misinformation make the human layer even more vulnerable.

The risks aren't only operational. **Cassandra Mack**, CISO at Tensorwave, highlighted the geopolitical stakes, noting that adversaries like China are investing heavily in AI, which creates systemic supply chain risks. **Sandip Wadje**, Global Head of Emerging Tech Risks at BNP Paribas, underscored the industry's weak baseline: ***"We ignored IT hygiene for years. Now we're playing catch-up at the very moment attackers are accelerating with AI."***

And behind it all, the economic asymmetry looms large. As **Amy Lemberger**, Senior Cybersecurity Consultant, put it: ***"They will always be a step ahead because they have more money and fewer constraints."***

Guidance: The Council's consensus is that attackers have seized an early advantage by adopting AI more quickly and with fewer guardrails. Defenders who hesitate will face a structural disadvantage. Adopting AI for triage, enrichment, and simulation is the only way to keep pace in an environment where the cost of attack is dropping toward zero.

What *Effective Defense* Looks Like in the AI Era

When the Council turned to the question of defense, a pattern emerged amongst the members that effective defense does not begin with novelty; it begins with fundamentals, executed faster, with more context, and supported by automation.

Antoinette Stevens, Principal Security Engineer at Ramp, captured the blunt truth: *"Do you have a shadow IT program? Are you patching quickly? It's the basics. You can add AI on top, but if you don't have the basics right (least privilege, vulnerability response, containment), you're still exposed."*

That focus on identity risk emerged repeatedly. **Jose Veitia**, Director of Information Security, emphasized that larger enterprises are beginning to assess their non-human identities and utilize automation to close gaps, but smaller organizations and local governments are still significantly behind. **Aunudrei Oliver**, Senior Director at Allianz Life, added that visibility and context are prerequisites for success, and without a disciplined SOC, AI has little to accelerate.

> "AI is valuable for reducing noise so teams can focus on higher-level capabilities, but you need the fundamentals in place. I look at it through the '6A' lens: Authentication, Authorization, Audit, plus Agents, Assets, and Automation. If those aren't working together, AI won't save you."
>
> — **Angelique Grado**

Others emphasized that fundamentals must be layered, not replaced. **Cassandra Mack**, CISO of Tensorwave, described layered defense across people, process, and technology as the most resilient model. **Peter Holcomb**, CEO of OptimoIT, agreed and added that automation is an ally, but only if it is governed like a sensitive asset itself, bound by least-privilege controls.

Even as AI brings speed, defenders warned about new blind spots. **Sean Todd**, CISO and AI Architect at Auditive, noted the shift in tempo: *"It's one thing to see pings at scale, it's another to correlate them into a probing pattern. Attackers are scanning with AI, so defenders must model context, not just alerts."* **Ryan Rosado**, Adjunct Faculty at Harvard, pointed to a lack of visibility into where AI models are being deployed and the need to segment development from production. Nonetheless, discerning between using AI in security operation defenses versus protecting the AI that exists within our networks. Much too often, AI is just a broad item from a security perspective, but these two could not be more of a juxtaposition; one is aiming to reduce the attack surface while the other is adding to it.

For many, the lesson was about resilience in the face of adversity on a large scale. **Susan Lloyd**, Director of Information Governance at iVisa, warned: *"You have to be ready for the fire hose. Attackers are using AI to create scale, and effective defense means building resilience into your processes before the wave hits."* And **Angelique Grado**, Principal and CISO, Vervation LLC, said resilience requires structure: *"AI is valuable for reducing noise so teams can focus on higher-level capabilities, but you need the fundamentals in place. I look at it through the '6A' lens: Authentication, Authorization, Audit, plus Agents, Assets, and Automation. If those aren't working together, AI won't save you."*

Guidance: The Council's perspective is that *"effective defense"* with AI is not about chasing the next tool or building the perfect detection model. It is about taking fundamentals like identity rigor, access controls, segmentation, and visibility, and applying AI to execute them faster, with more context, and at a greater scale. Basics plus AI assist beats novelty every time.

Humans and AI in the SOC

If the first two questions showed how AI is changing the tempo of conflict, the Council's discussion on the SOC made it clear that the future is not AI versus human, but AI plus human. The dividing line is labor. AI is well-suited to automate repetitive toil such as triage, correlation, enrichment, and reporting, while humans remain responsible for judgment, governance, and escalation. What emerged was simple but consistent: human-in-the-loop (HITL).

"…a human still reviews and closes the case. That's the balance that works right now."

— **Antoinette Stevens**

Practitioners described how this balance is already playing out. **Antoinette Stevens**, Principal Security Engineer at Ramp, explained that her team uses AI for L1 triage, but ***"a human still reviews and closes the case. That's the balance that works right now."*** **Amy Lemberger**, Senior Cybersecurity Consultant, cautioned against pushing beyond that balance: systems should never run in-line without checks and accountability.

Others warned of cultural risks if guardrails are ignored. **Ron Dilley**, Principal Architect at IS², argued that SOC platforms must resist the temptation to ***"just click the button,"*** reminding leaders that oversight is what prevents complacency. **AJ Debole**, Field CISO at Oracle, added that accountability can never be outsourced; humans must retain authority over command decisions, even as corporations seek efficiencies.

At the same time, members saw an opportunity for AI to take on the work that analysts dislike most. **Tomas Persson**, CISO at Omegapoint, pointed out that report writing and rote documentation are burdens that AI can easily lift, freeing humans for higher-value tasks. **Peter Holcomb**, CEO of OptimoIT, emphasized the forward-looking angle, noting AI's strength in predictive threat modeling, which is useful as long as it complements, not replaces, human modeling.

For some, the dividing line was organizational maturity. **Aunudrei Oliver**, Senior Director at Allianz Life, observed that AI can deliver measurable benefits only when a SOC is already process-driven and disciplined. And even among the most optimistic voices, guardrails remained non-negotiable. As **Susan Lloyd**, Director of Information Governance at iVisa, reminded her peers: ***"Computers cannot make a management decision. Analysts must act on the information."***

Guidance: The SOC is evolving into an AI-augmented, HITL-anchored model. AI will increasingly take on the burden of enrichment, correlation, and reporting, but judgment, escalation, and accountability will remain firmly human responsibilities. The organizations that strike a balance between automating toil and reinforcing human oversight will be the ones that adapt successfully to AI-driven threats.

"You have to treat these systems as if they possess the keys to the kingdom. If you wouldn't give an intern that kind of access, don't give it to an unbounded model."

— **Aunudrei Oliver**

New Risks from Autonomous Defense

If AI promises speed and scale, autonomy introduces a new set of risks. The Council discussed that handing over control without guardrails creates operational, ethical, and regulatory exposure that most organizations are not prepared to absorb. The risks range from catastrophic false positives to bias, privacy leakage, and the fundamental problem of AI's non-deterministic nature.

Several members warned of direct operational fallout. **Chris DeNoia**, CISO, put it plainly: ***"Catastrophic false positives are the biggest risk. If AI makes the wrong call in a fully automated system, the consequences can be immediate and severe."*** **Sean Todd**, CISO and AI Architect at Auditive, added that the problem runs deeper: ***"AI is probabilistic, while security requires near-perfect reliability. That mismatch, he argued, creates structural tension."***

Trust emerged as another fault line. **Ron Dilley**, Principal Architect at IS², noted that statistical and confirmation biases are already visible in AI systems and warned that a single public failure could provoke a regulatory backlash. **Peter Holcomb**, CEO of OptimoIT, highlighted related governance gaps, noting that data sovereignty and privacy leakage are live issues. His prescription was to keep humans in the loop for high-risk actions, backstop autonomy with audits, and apply just-in-time access controls to AI systems themselves.

Others pointed to the supply side of the problem. **Brandon Lindsay**, Director of Information Security at HIAS, argued that vendors are racing ahead without adequate safeguards, and that CISOs need to probe much harder into how training data is sourced and governed. **Cassandra Mack**, CISO at Tensorwave, brought it back to operations: ***"AI must be continuously monitored and treated as a crown jewel if it touches sensitive data."***

Not every Council voice saw autonomy as a new risk. **Amy Lemberger**, Senior Cybersecurity Consultant, offered a contrarian view: ***"I don't see additional risk, other than the risk of not using it, except when AI becomes your crown jewels. That's when your exposure multiplies."*** However, for others, legacy weaknesses were the greater concern. **Sandip Wadje**, Global Head of Emerging Technology Risks at BNP Paribas, pointed out that autonomy is often being layered onto environments riddled with poor hygiene and inconsistent standards, which is a fragile foundation for automation.

And at the heart of it all, **Aunudrei Oliver**, Senior Director at Allianz Life, returned to governance: ***"You have to treat these systems as if they possess the keys to the kingdom. If you wouldn't give an intern that kind of access, don't give it to an unbounded model."***

Guidance: The Council agreed that autonomy should be introduced cautiously and governed rigorously. That means human-in-the-loop for sensitive actions, role-based and just-in-time access controls for AI systems, continuous audits, and strict data minimization. AI may accelerate defense, but without strong guardrails, it risks introducing as much exposure as it prevents.

Board-Level Priorities

When the conversation turned to the boardroom, the Council's message was consistent around not chasing AI as a shiny object. Boards want awareness, ROI, and clarity on risk. Practitioners emphasized that the right place to start is education and hygiene, building a workforce that understands AI's impact, and mitigating identity and data risks before expanding into automation with measurable outcomes.

> "AI will eventually enable super-lean security teams to do the work of many. But that only works if the board and executive team invest in building the right foundation now."
>
> — **Adnan Dakhwe**

Several members described a phased approach. **Chris DeNoia**, CISO and author, advised sequencing investments: ***"Start with education, then visibility, then enrichment, and finally some automated response. You have to sequence it or the investment won't stick."*** **Susan Lloyd**, Director of Information Governance at iVisa, echoed that sentiment, warning that boards are asking the wrong questions if they leap to tools before fixing fundamentals.

Identity and data governance were repeatedly highlighted as the real starting line. **Jose Veitia**, Director of Information Security, framed it as a three-pronged priority: ***"identity risk, data risk, and third-party risk."*** **Angelique Grado**, Director of Information Security at Gamechanger, tied that back to growth, noting that boards will listen when security is positioned as protecting opportunity, not just preventing loss.

Measurement was another recurring theme. **Ryan Rosado**, an Adjunct Faculty Member at Harvard, urged leaders to define problems clearly and demonstrate progress through metrics. She suggested that AI security and AI efficiency should be tracked as separate but complementary goals.

Education, however, was seen as more than training modules. **Tomas Persson**, CISO at Omegapoint, referred to it as cultural readiness, emphasizing that key employees must be educated above all else. **Cassandra Mack**, CISO at Tensorwave, warned that training often falls by the wayside unless leaders actively fight for it at the board level. **Amy Lemberger**, Senior Cybersecurity Consultant, sharpened the point, urging security leaders to hold directors accountable: ***"if boards cannot articulate the business objectives for AI, funding requests risk being wasted."***

Looking ahead, **Adnan Dakhwe**, CISO at DelphinusCyber, envisions exponential efficiency gains as a long-term outcome: ***"AI will eventually enable super-lean security teams to do the work of many. But that only works if the board and executive team invest in building the right foundation now."***

Guidance: The Council's advice to boards starts with awareness and training, then addresses identity, data, and third-party risk. Invest in automation only after establishing success criteria and developing metrics to measure ROI. Boards that demand clarity before spending, and security leaders who push for it, will avoid both waste and unnecessary exposure.

Reflections

Across two workshops, it became clear that AI is not looked upon as a silver bullet. It doesn't guarantee security, but it does change the tempo of conflict. In the hands of attackers, it lowers costs and expands reach. In the hands of defenders, it can compress timelines, reduce toil, and make small teams more effective. The outcome depends less on the technology itself and more on how deliberately organizations utilize it.

The Council's collective answer to the question of **"effective defense"** was fundamentals executed at speed. Hygiene, access controls, and segmentation are still the bedrock of security. What AI brings is the ability to carry them out faster, with more context, and at scale. As **Jose Veitia** put it, the real battle lines are around identity and access. As **Antoinette Stevens** reminded the group, none of this works without getting the basics right first.

On the future of the SOC, the consensus was equally firm: AI can take on toil, but humans must govern risk. Human-in-the-loop is no longer a principle; it is doctrine. Automation should enrich, correlate, and report, but judgment and escalation remain human responsibilities.

The risks of autonomy (bias, privacy leakage, catastrophic false positives, non-determinism) are real, and they must be managed with audits, RBAC, and strong guardrails. As **Chris DeNoia** warned, one bad automation can break production overnight.

And when it comes to the boardroom, the Council's message was simple — **start with awareness and guardrails**. Invest in training, identity, and data protection before leaping to automation. Only then should boards fund AI-driven defenses, and even then, only with clear success criteria and metrics for ROI.

This white paper is not a theoretical view of the future. It is a record of what practitioners are seeing and doing right now, in the early stages of AI's impact on cyber defense. Their guidance is both pragmatic and urgent: adopt AI where it accelerates fundamentals, govern it with human oversight, and build the culture and awareness to use it responsibly. Anything less risks ceding permanent advantage to attackers who are already moving faster.

Join the AI Security Council

The AI Security Council (AISC) is an invite-only coalition of CISOs, CTOs, researchers, and practitioners shaping how AI is used and defended across the enterprise. It's where security leaders anticipate adversarial AI, share field-tested strategies, and publish actionable frameworks for defending modern environments.

If you want a seat at the table shaping how security evolves with AI, **apply here**.

https://www.tuskira.ai/ai-security-council

AI SECURITY COUNCIL